

ALCF Resources

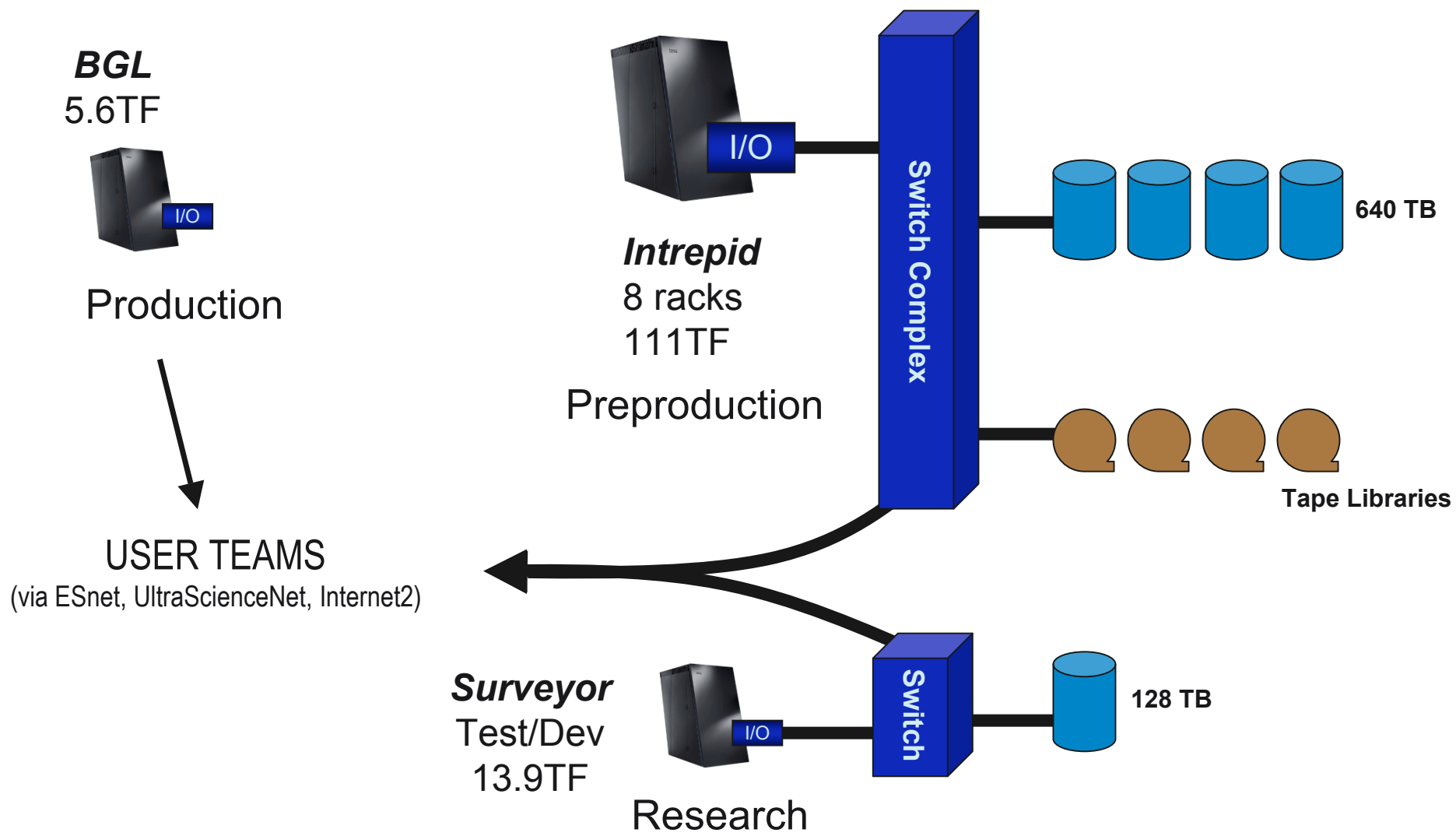
Susan Coghlan
Deputy Director



Argonne National Laboratory is managed by
The University of Chicago for the U.S. Department of Energy

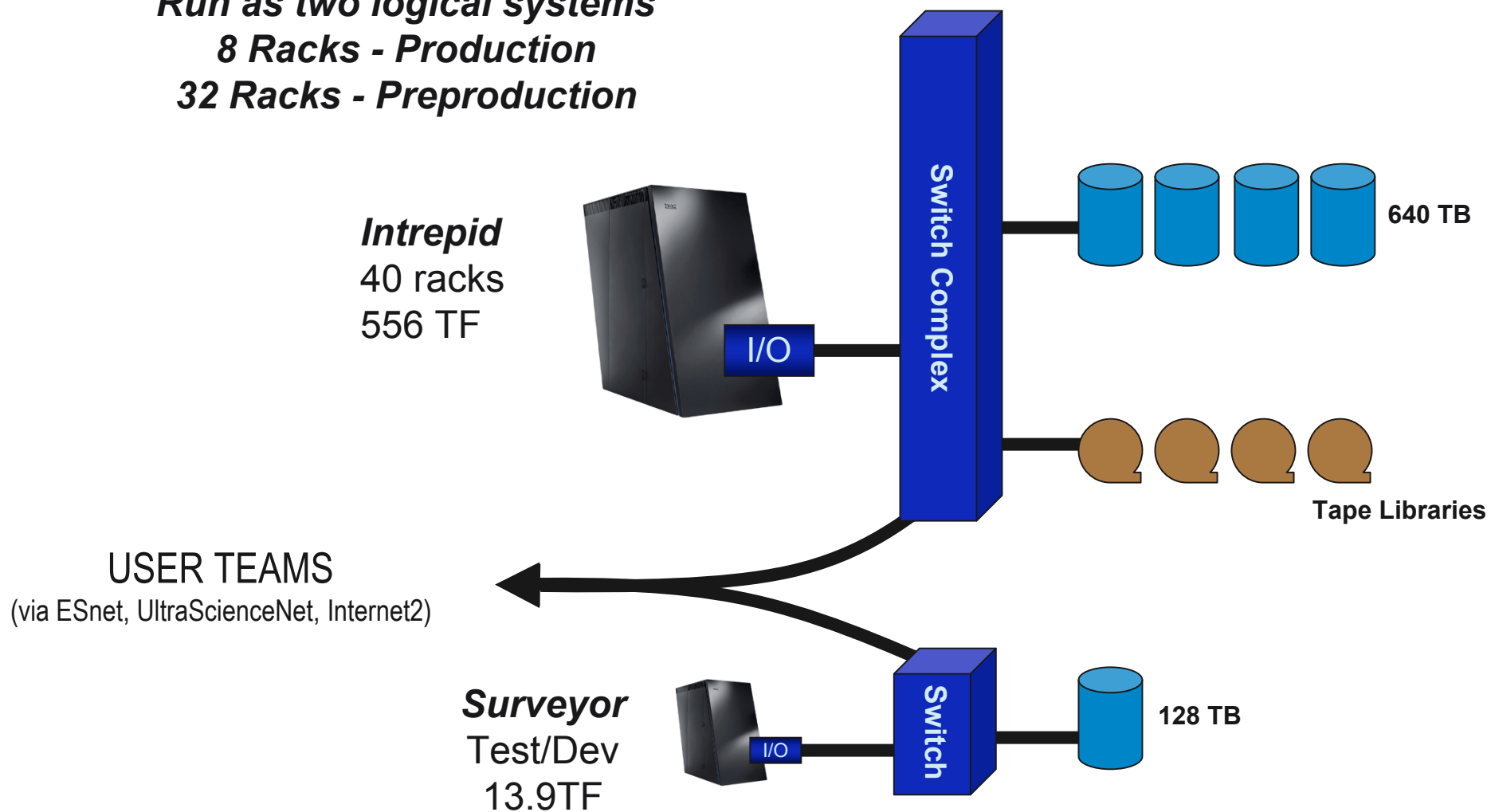


ALCF Resources - Current

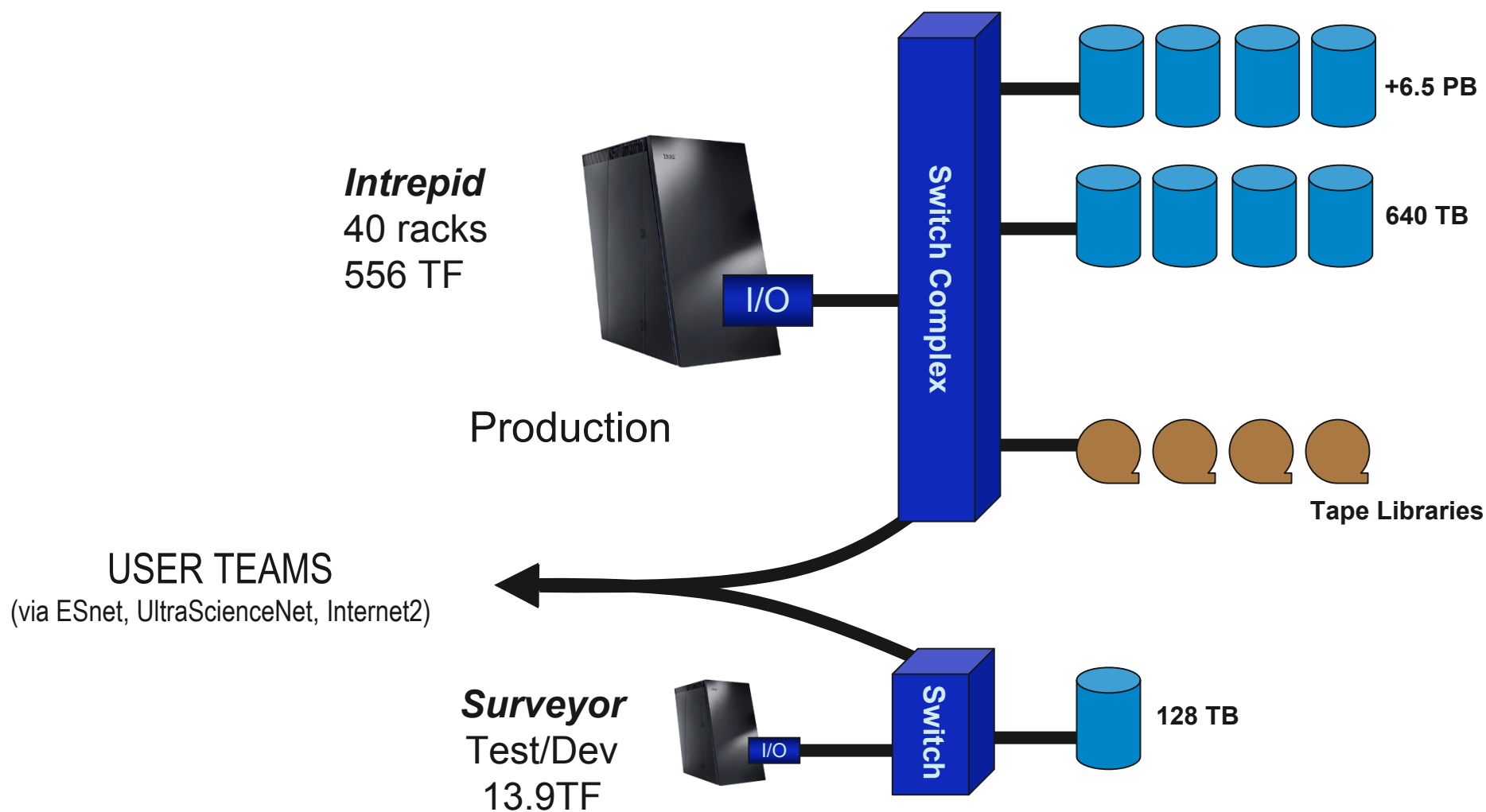


ALCF Resources - Later this year

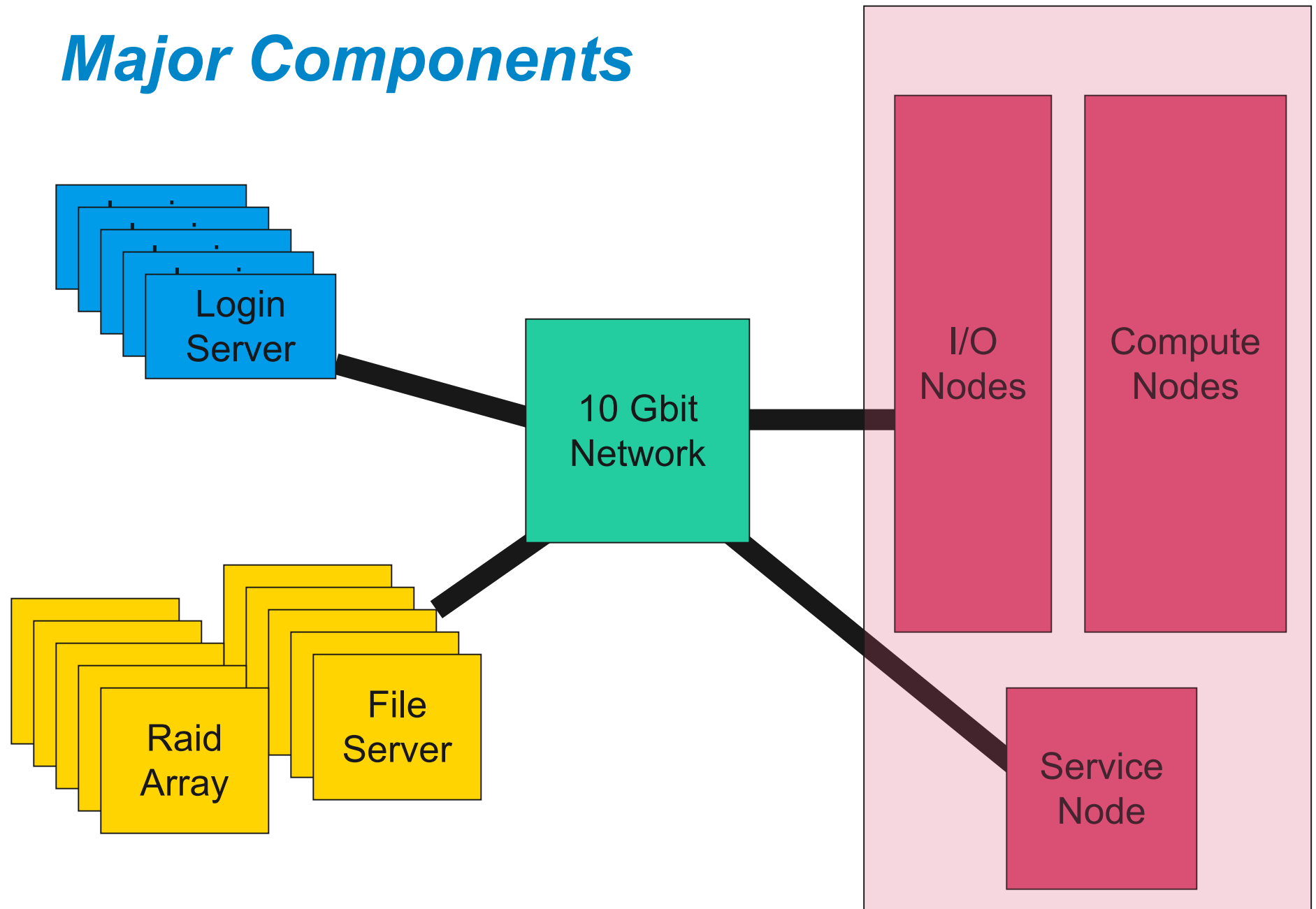
Run as two logical systems
8 Racks - Production
32 Racks - Preproduction



ALCF Resources - 2009



Major Components



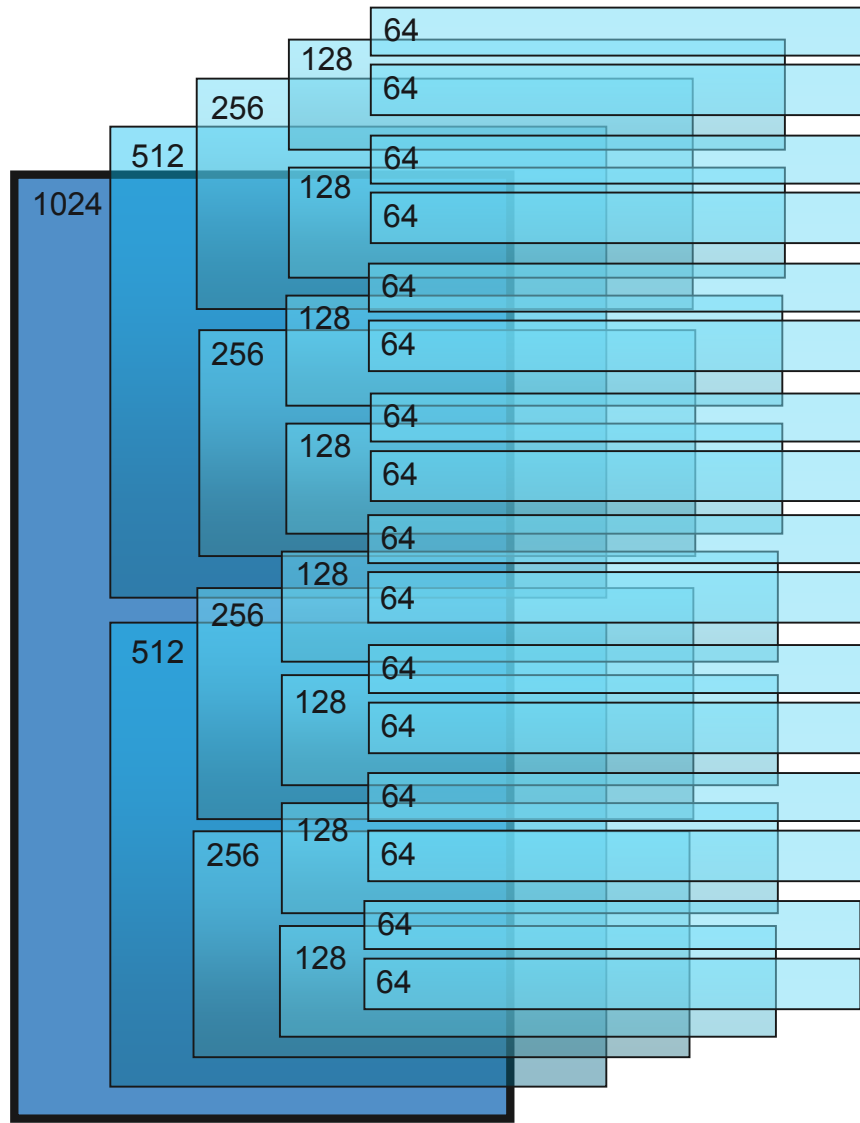
Configuration Details

System	Surveyor	Intrepid (8 racks)
Function	Test & Development	Production INCITE
Login access	surveyor.alcf.anl.gov	intrepid.alcf.anl.gov
# Nodes/# Cores	1024 / 4096	8192 / 32768
Memory (2GB per node)	2TB	16TB
# I/O nodes (1/64 ratio)	16	128
# File servers	4 Home/Data (Shared)	16 Data, 8 Home

- Compile and submit jobs on the login servers
 - Multiple servers, DNS round robin
 - Login[1-x].<machinename>.alcf.anl.gov
- No direct access to compute nodes

BG Partitions (“blocks”)

Single Rack



- 1 I/O node for each 64 compute nodes, hardwired to specific set of 64
 - *Minimum partition size of 64 nodes*
- Partition sizes: 64, 128, 256, 512, 1024
 - *Any partition < 512 nodes will get a mesh network layout and not a torus.*
 - *Any partition < 512 nodes will get a non-optimal I/O tree network.*
 - *Do not do performance testing on < 512 nodes*
- Smaller partitions are enclosed inside of larger ones
 - *Not all partitions are available at all times*
 - *Once a job is running on one of the smaller partitions, no jobs can run on the enclosing larger partitions*
- Configuration changes frequently
 - ***partlist*** shows partition state
- Processes are spread out in a pre-defined mapping, alternate and sophisticated mappings are possible

File Systems - Surveyor

File System	Home	Data	Local
Mount pt	/home	/pvfs-surveyor	/scratch
Size	15TB	88TB	69GB each
Allocation	10GB	100GB	-
FS Type	GPFS	PVFS	XFS
Backed Up?	Snapshots	No	No
Visible on BG/P	Yes	Yes	No
Uses	General use, robust, reliable.	High speed I/O, large file, parallel file access.	Reasonably fast, local, temporary.
Warnings	Limited Space.	Binaries cannot be executed.	Login servers only, not cross mounted.

File Systems - Intrepid (Current)

File System	Home	Data	Local
Mount pt	/home	/gpfs1	/scratch
Size	59TB	233TB	69GB each
FS Type	GPFS	GPFS	XFS
Backed Up?	Snapshots	No	No
Visible on BG/P	Yes	Yes	No
Uses	General use, robust, reliable.	Medium speed I/O, large file, parallel file access.	Reasonably fast, local, temporary .
Warnings	Limited Space.	Still in testing, expect changes.	Login servers only, not cross mounted.



Resource Mgr and Job Scheduler

- Cobalt - locally developed
- Standard commands (qsub, qstat, qdel, qalter)
- **Scheduling policy is in flux, expect changes**
- Surveyor queues (FIFO based, with some exceptions)
 - **default**: routes to appropriate queue
 - **short**: $\leq 30\text{min}$ and < 1024 nodes
 - Use 'qstat -Q' and 'partlist' commands
 - **R.workshop** - use for all of your jobs today and tomorrow
 - Special purpose queues, generally for specific people and needs
- Reservations
 - Required for anything larger than 512 nodes
 - Email reservation requests to ***support@alcf.anl.gov***
 - View reservations with 'showres'
- Standing reservations
 - Preventative maintenance reservation: Each Monday, 8am to 8am
 - Scaling run day: Each Tuesday 9am-5pm CST (upon request)

Backups and Archival

- Snapshots of home directories are done nightly
~/.snapshot
- Backups of home directories to tape are coming soon
- Data directories will not be backed up
- Archival will be through HPSS
- HPSS planned for April time frame
- gridFTP will be available for external access to the file servers
- gridFTP will also be used to stream data direct to tape servers

Allocation Management

- Allocations are different for different resources
 - Surveyor: small, startup allocations
 - Intrepid: official INCITE allocations
- Charges are based on the partition size, NOT the # of nodes or cores used!
- Reservations are charged for the full time they are active
- Will be managed with clusterbase
 - Use the 'cbank' command (see 'cbank --help')
- Examples:
 - # list all charges against a particular project
 - *cbank -l charge -p <projectname>*
 - # list all active allocations for a particular project
 - *cbank -l allocation -p <projectname>*
- If you don't know your project name
 - `cat /etc/projects.map | grep <username>`

Mailing Lists

- For each resource there are two mailing lists
- users@<resource>.alcf.anl.gov
 - Mandatory, auto-built from all users with active accounts
 - Important announcements impacting the entire community
 - *Security issues*
 - *Major downtimes*
 - *Policy changes*
 - *Long-term news*
- notify@<resource>.alcf.anl.gov
 - For the active community
 - Operational status announcements
 - Initially subscribed with account creation
 - Subscribe/unsubscribe as you wish
 - *Use the command 'notifyme'*

Cyber Security

- Argonne computer user agreements
 - Agreed to at account request time
- Standard Argonne computer security rules apply
 - No sharing accounts
 - Acceptable use
 - Etc.
- No passwords are allowed for accessing the systems
 - SSH keys used to access Surveyor
 - CRYPTOCARD token required to access Intrepid after production
 - *CRYPTOCARD tokens will work for Surveyor as well*
- Data policies are available on the web:
 - <http://www.alcf.anl.gov/support/usingALCF/docs/dataprivacy.php>
 - If you have prohibited data (PII, UNCI, etc.) please contact support@alcf.anl.gov

Help!

- Problems or questions

email: support@alcf.anl.gov

phone: 630-252-3111 (866-508-9181)

- Resources:

- ALCF web pages: <http://www.alcf.anl.gov>

- Good starter page:

- <http://www.alcf.anl.gov/support/usingALCF/usingsystem.php>
(contains link to online IBM BG redbooks)